Cooperation Studies of Catastrophe Avoidance

Implications For Climate Change Negotiations

PSYC370: The Psychology of Cooperation
Mark Hurlstone

“We see the world as it is but not how it may have been”

The Counterfactual Problem

Background Reading

Hurlstone M. J. et al. (2017). Cooperation studies of catastrophe avoidance: Implications for climate negotiations. *Climatic Change, 140,* 119-133

Climatic Change
DOI 10.1007/s10584-016-1838-3

REVIEW ARTICLE

Cooperation studies of catastrophe avoidance: implications for climate negotiations

Mark J. Hurlstone[1,2] · Susie Wang[1,2] · Annabel Price[1] · Zoe Leviston[1,2] · Iain Walker[1,2,3]

Received: 27 October 2015 / Accepted: 8 October 2016

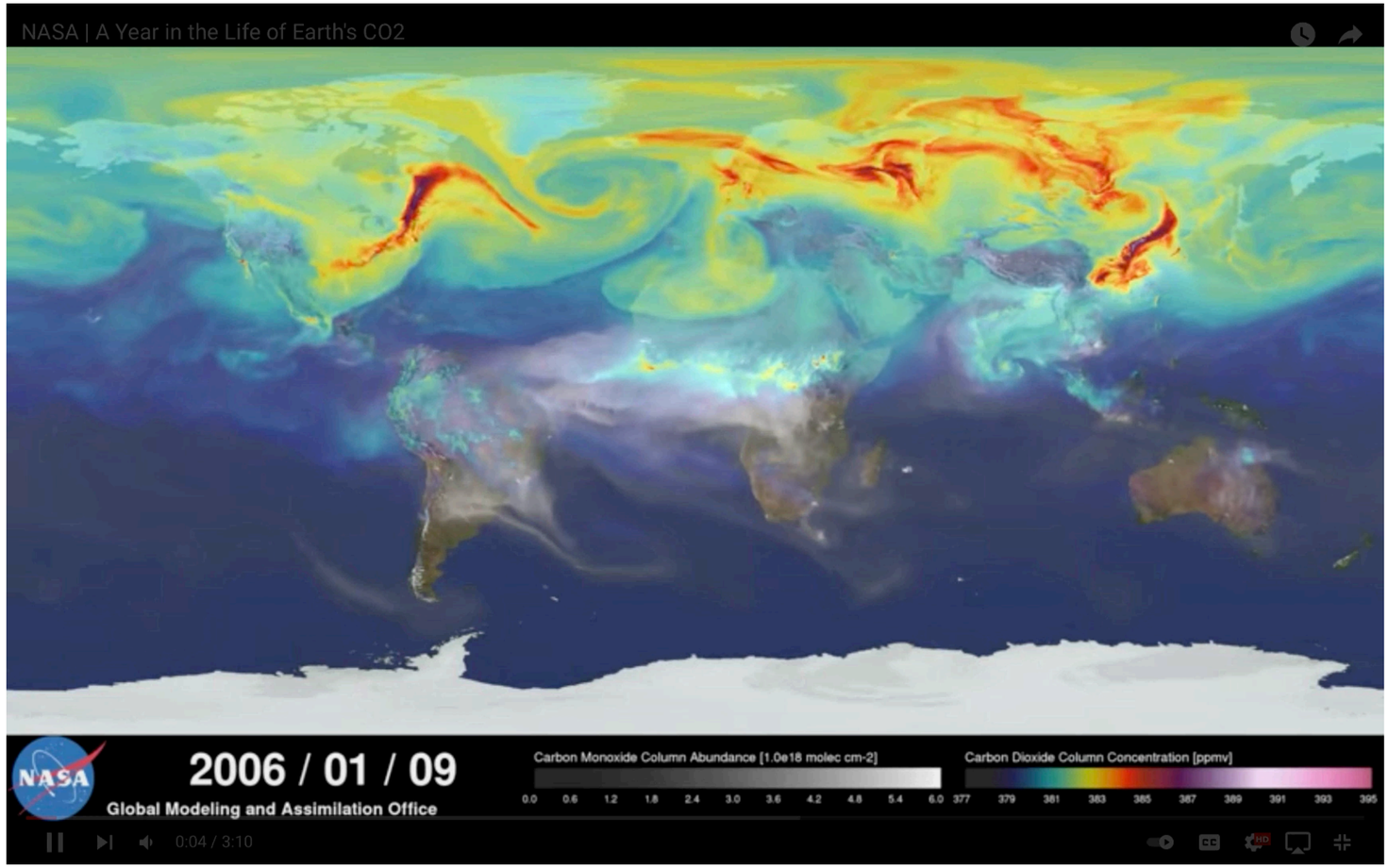
Abstract The landmark agreement recently negotiated in Paris represents an ambitious plan to combat climate change. Nevertheless, countries' current climate pledges are insufficient to achieve the agreement's goal of keeping global mean temperature rise "well below" 2 °C. It is apparent that climate negotiators need to be equipped with additional strategies for fostering cooperation if a climate catastrophe is to be averted. We review the results arising from an emerging literature in which the problem of avoiding dangerous climate change has been simulated using cooperation experiments in which individuals play a game requiring collective action to avert a catastrophe. This literature has uncovered five key variables that influence the likelihood of avoiding disaster: (1) the perceived risk of collective failure, (2) inequalities in historical responsibility, wealth, and risk exposure, (3) uncertainty surrounding the threshold for catastrophe, (4) intergenerational discounting, and (5) the prospect of reward or punishment based on reputation. Along with the results of a recent experimental assessment of the key instruments of the Paris Agreement, we consider how knowledge of the effects of these variables might be harnessed by climate negotiators to improve the prospects of reaching a solution to global climate change.

Keywords Climate change · Cooperation · Climate negotiations · Global public good · Collective-risk social dilemma

A Most Dangerous Game

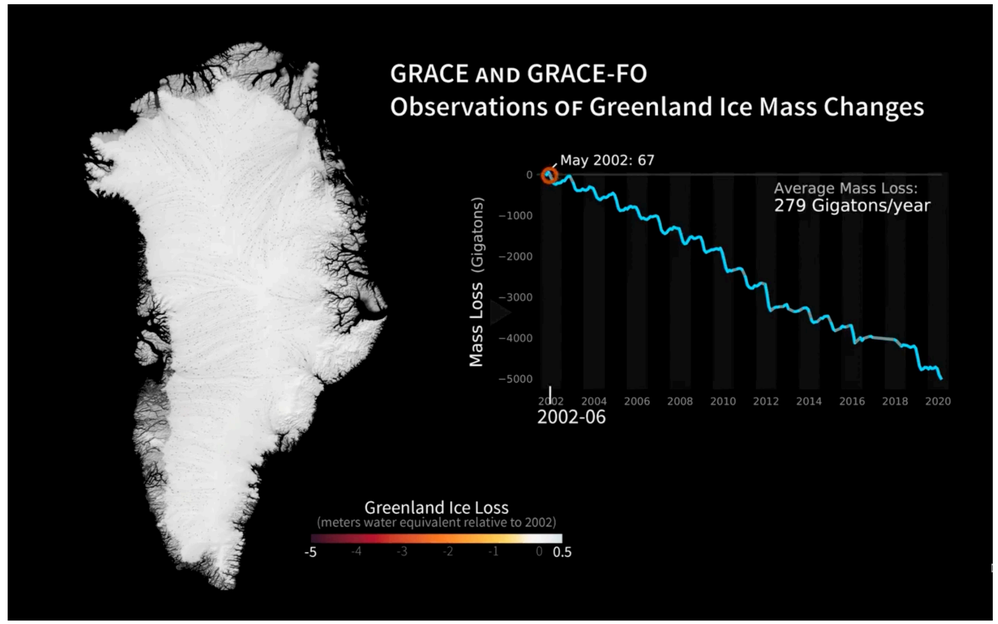
Climate Change

- Mean global surface temperature has already increased by **1.2°C** since pre-industrial times

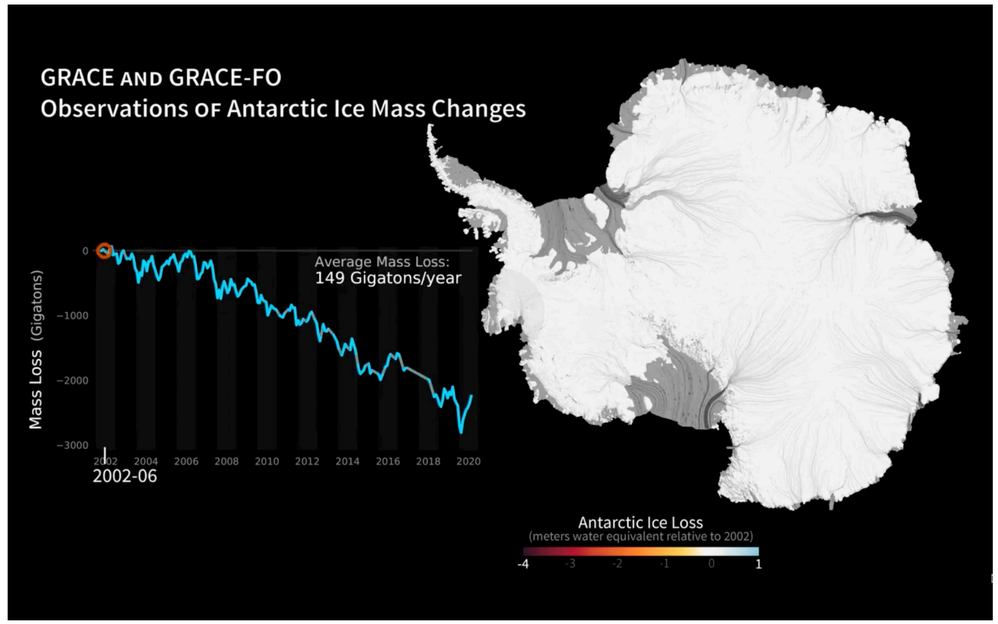


Gradual vs. Dangerous Climate Change

- **Gradual climate change**
  - *Linear* process
  - Emissions increase a little → concentrations increase a little → mean global temperature increases a little → human impacts increase a little
- **Dangerous climate change**
  - *Nonlinear* process
  - Emissions increase a little → concentrations increase a little → mean global temperature increases a little → human impacts, near a threshold, increase ***a lot***



Greenland Ice Sheet Retreat



West Antarctic Ice Sheet Retreat

Climate Change: The Global Public Good

- Protection of the Earth's climate is a **global public good**
- Two key properties:
  - Each country benefits the same way from a safe climate (**non-rivalry**)
  - No country can be excluded from climate protection (**non-excludability**)
- These properties of the public good leave it vulnerable to **free riding**

A Brief History of The Climate Change Negotiations

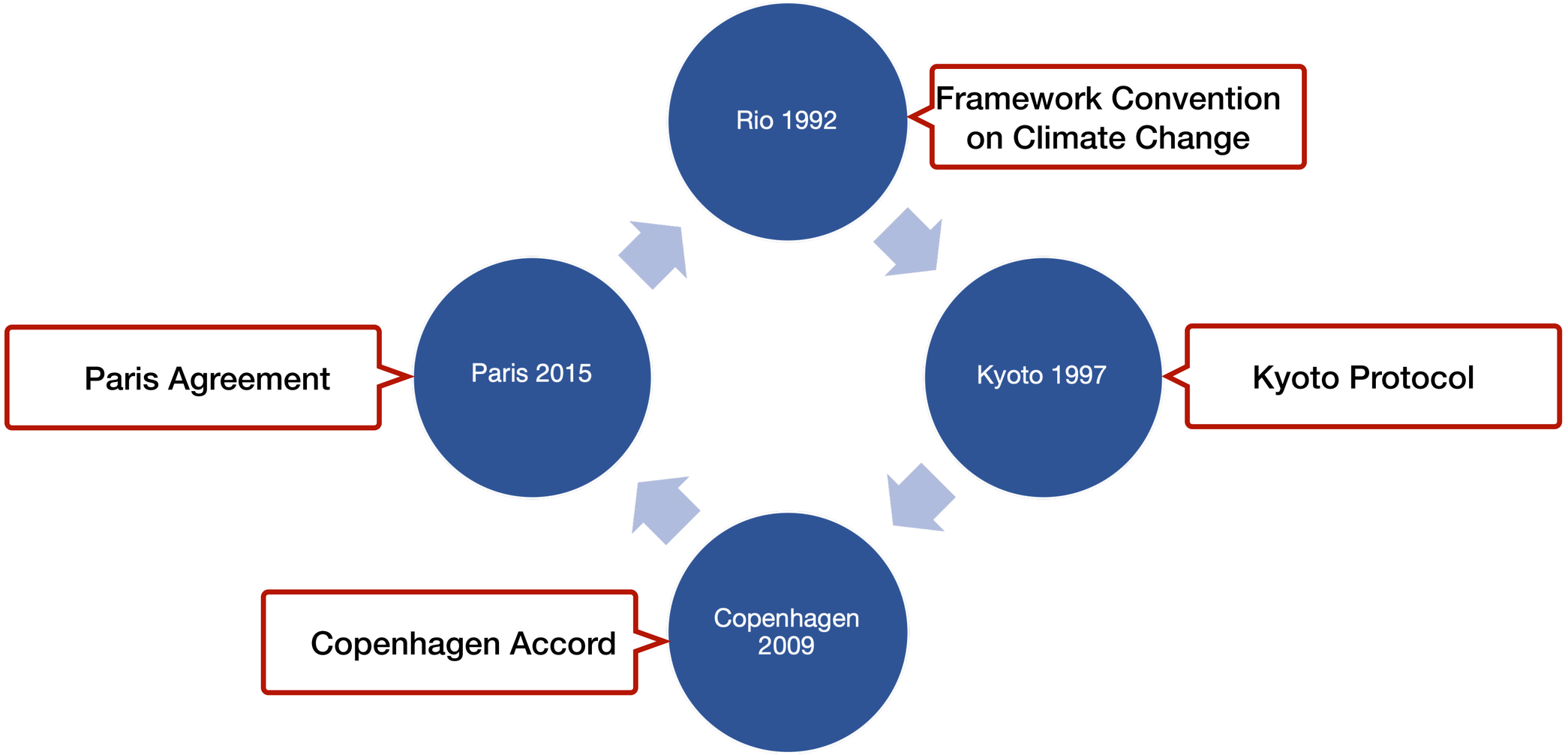
The Challenge of Creating An Effective Climate Treaty

Treaties As Instruments of “Strategy”

- The climate problem is difficult to solve due to the principle of **sovereignty**
- Requires **voluntary cooperation**
- **Climate treaties** are the instruments for achieving this
- To sustain cooperation, a treaty must deter non-participation and non-compliance—it must **strategically transform the incentive system**
- Creating a treaty that tells parties to reduce their emissions is easy—getting them to **participate** and **comply** is harder

Climate Negotiations

Major Milestones



Climate Negotiations

Rio 1992

- Establishment of United Nations Framework Convention on Climate Change (UNFCCC)
- “stabilise atmospheric concentrations in the atmosphere at a level that would prevent *dangerous* [emphasis added] anthropogenic interference with the climate system” (Article 1, UNFCCC; 1992)

Climate Negotiations

Kyoto 1997

- Kyoto protocol established
- “Top-down approach” to climate diplomacy based on targets and timetables
- Legally binding but easy to opt out (as Canada did)
- No enforcement mechanism
- No threshold for dangerous interference

Climate Negotiations

Copenhagen 2009

- Copenhagen Accord
- Identified a threshold for “dangerous anthropogenic interference”
- “in accordance with the *scientific view* [emphasis added] global temperature should be below 2°C (Article 1, UNFCCC, 2009)
- Transition to bottom-up “pledge-and-review” approach to climate diplomacy

Climate Negotiations

Paris 2015

- Paris Agreement
- Retains the 2-degree-target of the Copenhagen Accord
- Recognises the importance of **pursuing 1.5°C**—Article 2.1 includes the phrase “well below 2 degrees Celsius”
- Paris pledges are **only enough to prevent more than 2.7°C** of warming

The Game We Cannot Afford To Lose

Cooperation Studies

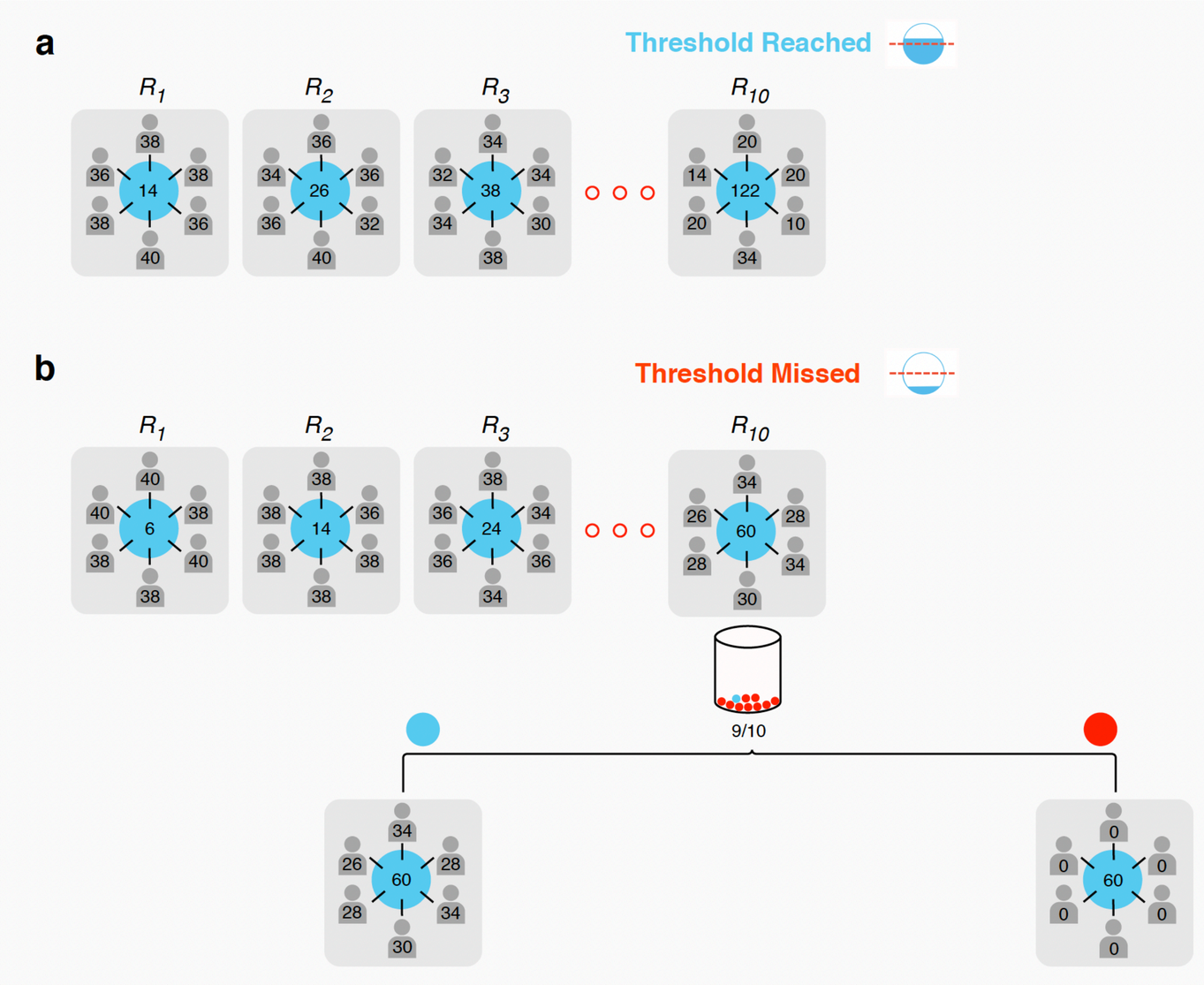
- Cooperation studies **simulate the problem of avoiding dangerous climate change** in the laboratory
- Groups of individuals play a game that must be solved through cooperation
- Can identify when people are likely to **cooperate** and when they will **stubbornly refuse**
- Provides insights into **which strategies** are likely to **succeed** in climate negotiations

A Climate Catastrophe Avoidance Game

Collective-Risk Social Dilemma

Milinski et al. (2008)

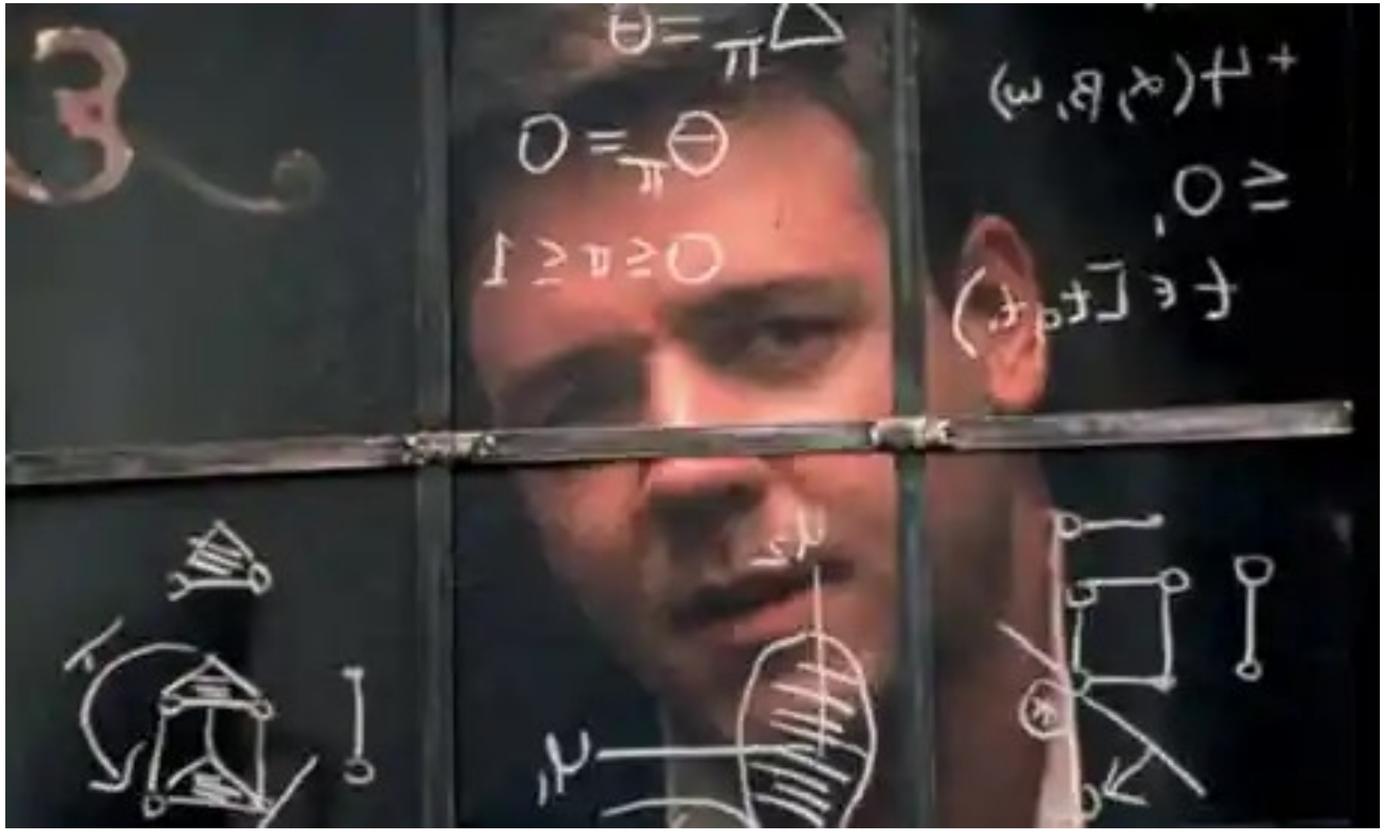
- Groups of six players
- Each given a €40 operating fund
- On each of 10 rounds, each player decides whether to invest €0, €2, or €4
- If total investments ≥ €120 by the end of the game, catastrophe is averted
- Otherwise catastrophe occurs with a pre-specified probability (e.g., 90%)



Collective-Risk Social Dilemma

Equilibria

- **Coordination game**
- Two “Nash” equilibria:
  - **Safe equilibrium:** all players contribute €20 and catastrophe is averted
  - **Dangerous equilibrium:** all players contribute €0 and catastrophe occurs
- The safe equilibrium is “focal” (Schelling, 1960)



Russell Crowe as John Forbes Nash in the movie “A Beautiful Mind”

Factors Affecting Cooperation

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Factors Affecting Cooperation

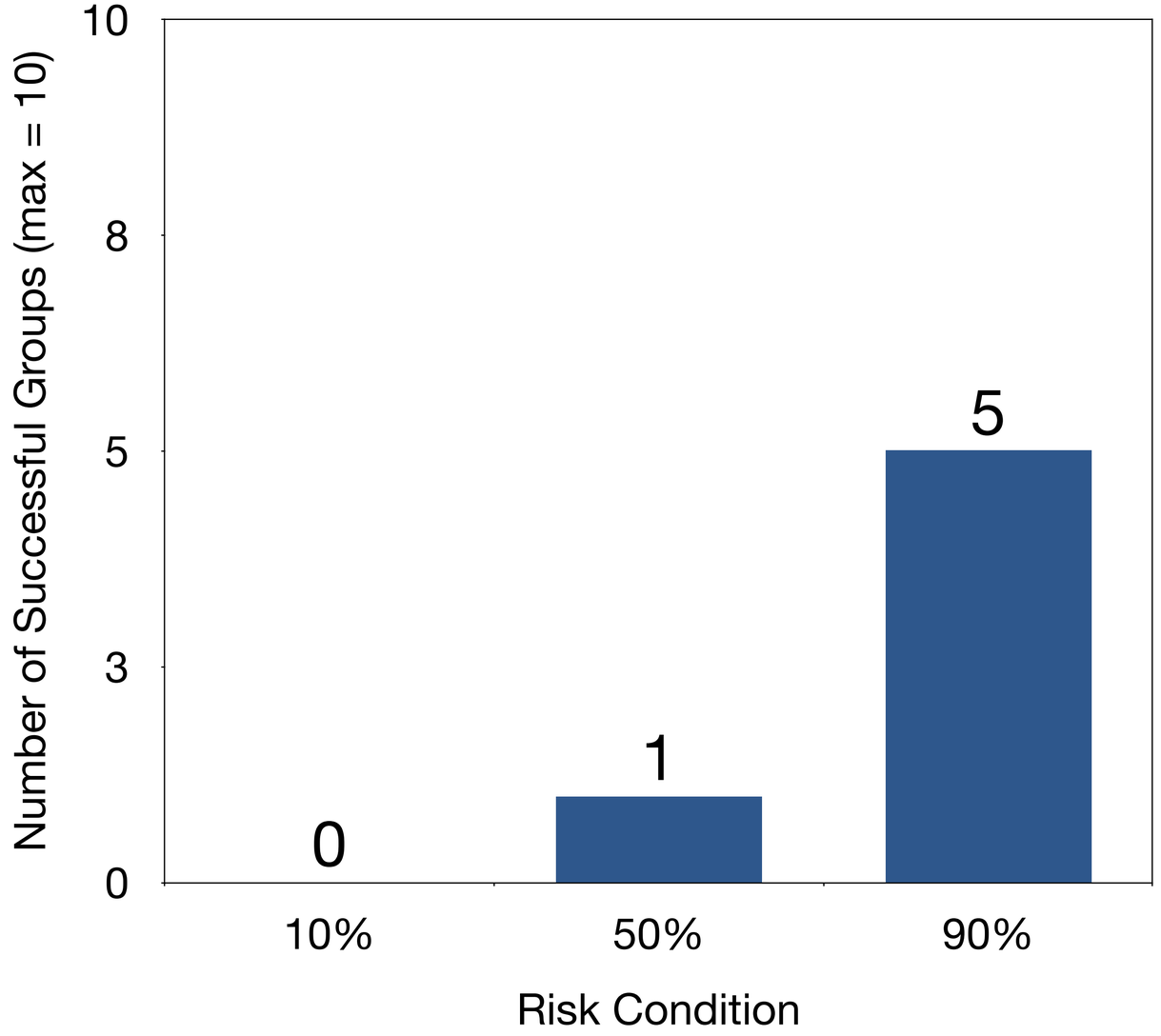
Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Perception of Risk

Milinski et al. (2008)

- Manipulated the probability of catastrophe occurring
- Three risk conditions:
  - 10 % (low risk)
  - 50 % (moderate risk)
  - 90 % (high risk)



Implications For Climate Negotiations

Risk Communication

- Climate negotiators rely on assessment reports of the Intergovernmental Panel on Climate Change (IPCC), which uses probabilistic statements (e.g., “Very Likely”) to communicate risks

Likelihood Scale of the Intergovernmental Panel on Climate Change

Phrase	Likelihood of occurrence
Virtually certain	> 99%
Very likely	> 90%
Likely	> 66%
About as likely as not	33% to 66%
Unlikely	< 33%
Very unlikely	< 10%
Exceptionally unlikely	< 1%

Implications For Climate Negotiations

Risk Communication

- People **underestimate** the verbal IPCC statements (Budescu et al., 2012, 2014):
  - “it is ***very likely*** that hot extremes, heat waves, and heavy precipitation events will continue to become more frequent”
  - 25% of respondents had in mind probabilities lower than 70%!
- The tendency can be reduced by using a verbal and numerical scale:
  - “it is ***very likely*** **(>90%)** that hot extremes, heat waves, and heavy precipitation events will continue to become more frequent”
- The IPCC has been encouraged to move to this “dual-scale” scale format

Implications For Climate Negotiations

Risk Communication

- The scientific community may have **understated** climate risks
- Evidence of systematic under-predictions by climate scientists of key attributes of climate change (Brysse et al. 2013) and undue conservatism regarding climate risks in assessment reports of the IPCC
- Lewandowsky et al. (2015) attribute this tendency to “err on the side of least drama” to “seepage”—the influence of contrarian talking points from public discourse about climate change into scientific thinking
- Climate negotiators and IPCC need methods to detect and avoid seepage

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Inequality

- In the collective-risk game considered so far, all players are equal
- This is not an accurate reflection of the real climate change game where inequalities exist in terms of:
  - Historic responsibility
  - Wealth
  - Risk exposure

Historic Responsibility and Wealth

Tavoni et al. (2011)

- Collective-risk social dilemma divided into “passive” and “active” phase sub-components, with a 50 % risk of catastrophe:
- **Passive phase:** Rounds 1—3
  - *Equal condition*: all six players forced to contribute €2 per round (all players left this phase with €34 each)
  - *Unequal condition*: half the players forced to contribute €4 per round, whereas the other half were forced to contribute €0 (“poor players” left this phase with €28; “rich players” with €40)
- **Active phase:** Rounds 4–10
  - Players decided for themselves how much to invest per round

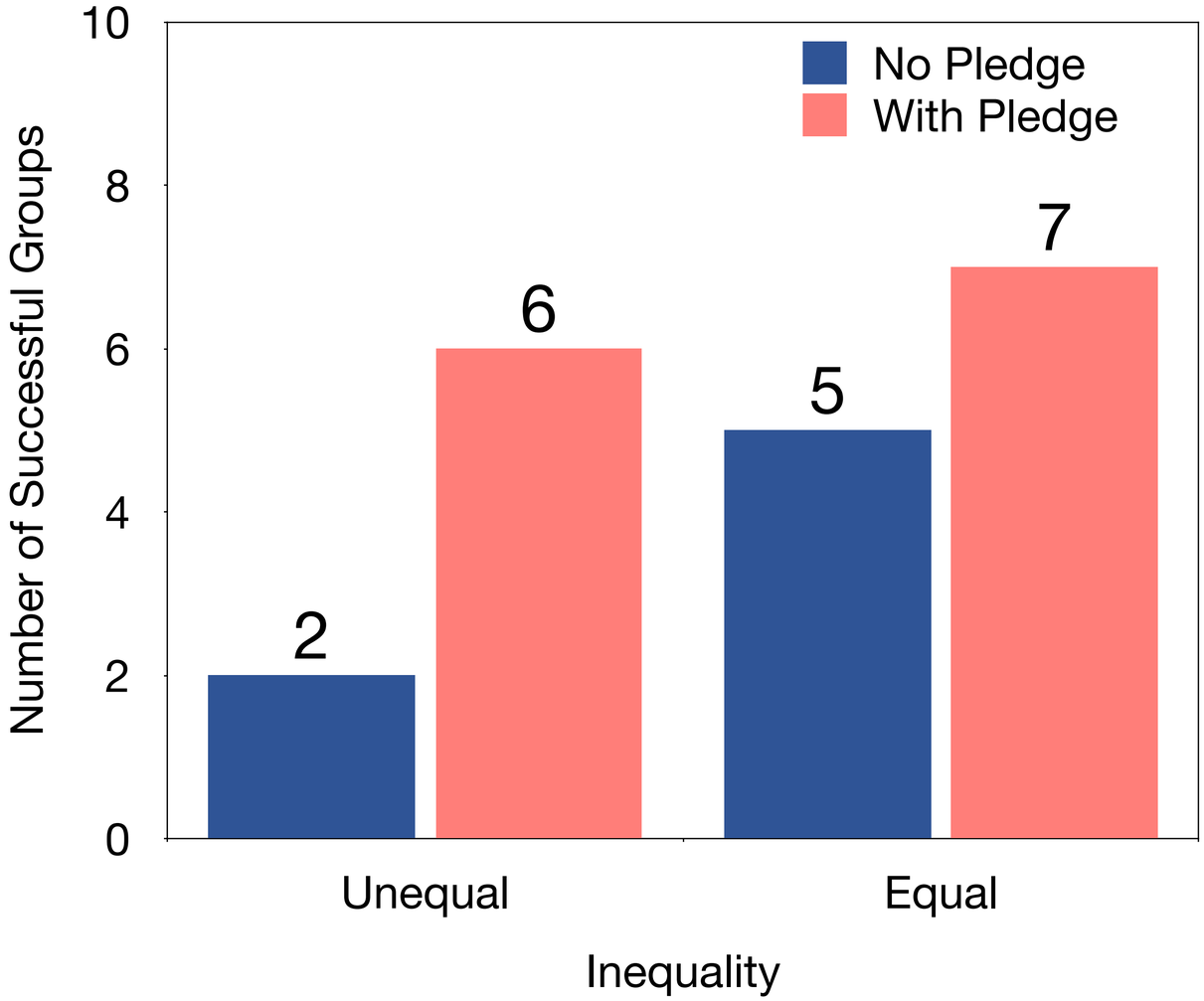
Historic Responsibility and Wealth

Tavoni et al. (2011)

- In addition to the equality manipulation, Tavoni et al. used a communication manipulation
  - *With-pledge condition*: at the end of rounds 3 and 7, players announced how much they intended to invest over the next three rounds by submitting non-binding pledges
  - *No-pledge condition*: communication between players was prevented
- This coordination device was designed to mimic the “pledge-and-review” instrument in the Copenhagen Accord/Paris Agreement

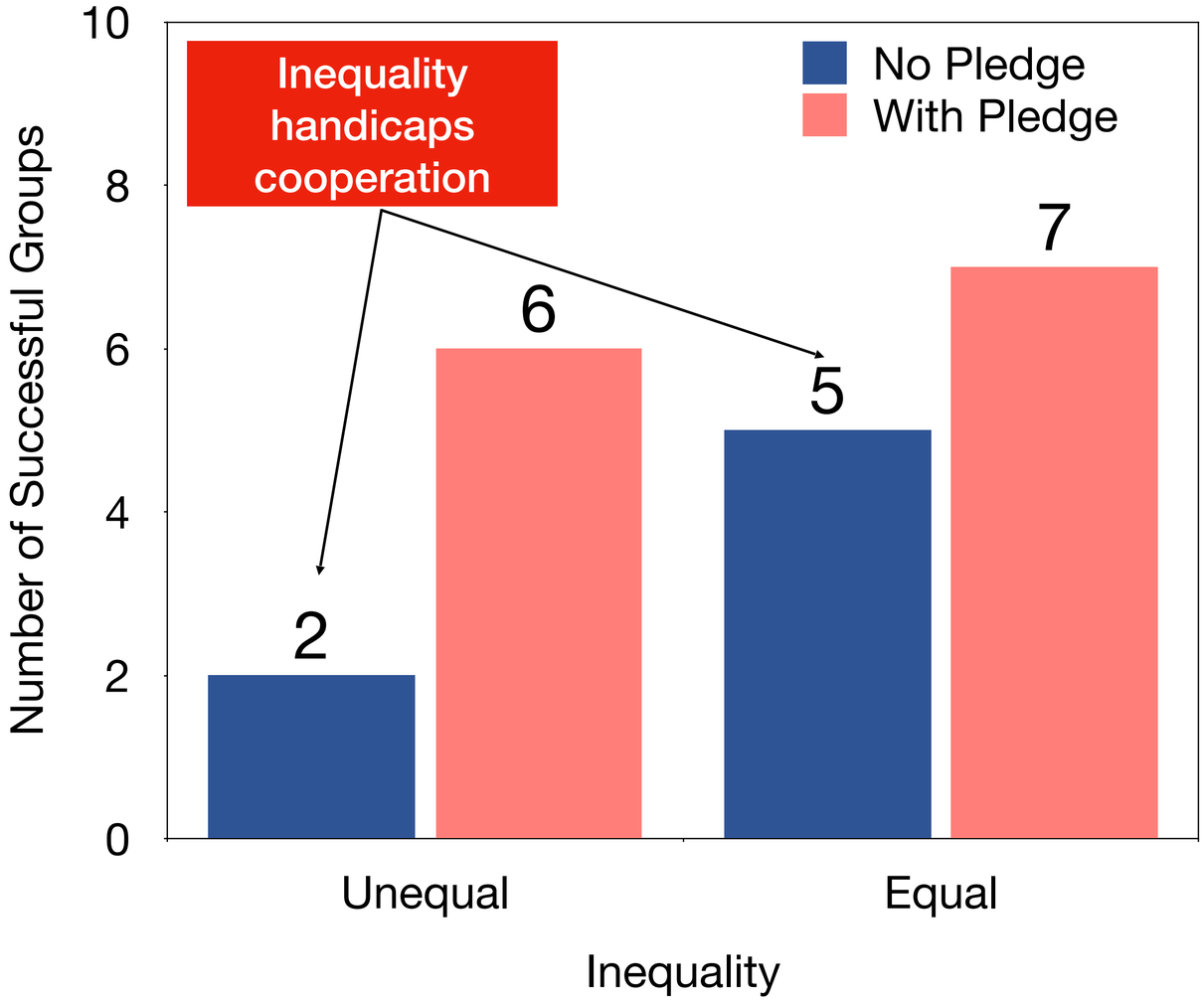
Historic Responsibility and Wealth

Tavoni et al. (2011)



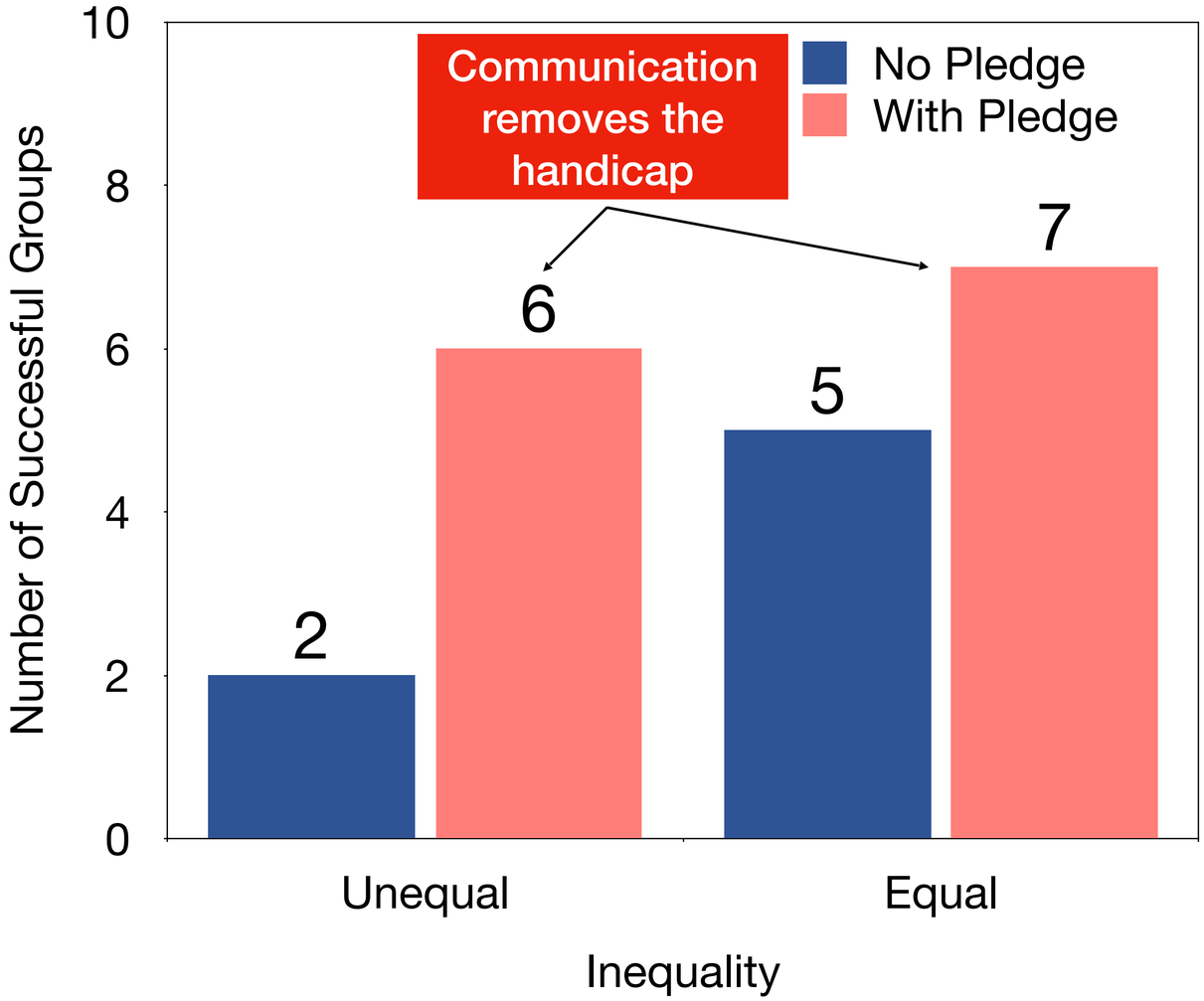
Historic Responsibility and Wealth

Tavoni et al. (2011)



Historic Responsibility and Wealth

Tavoni et al. (2011)



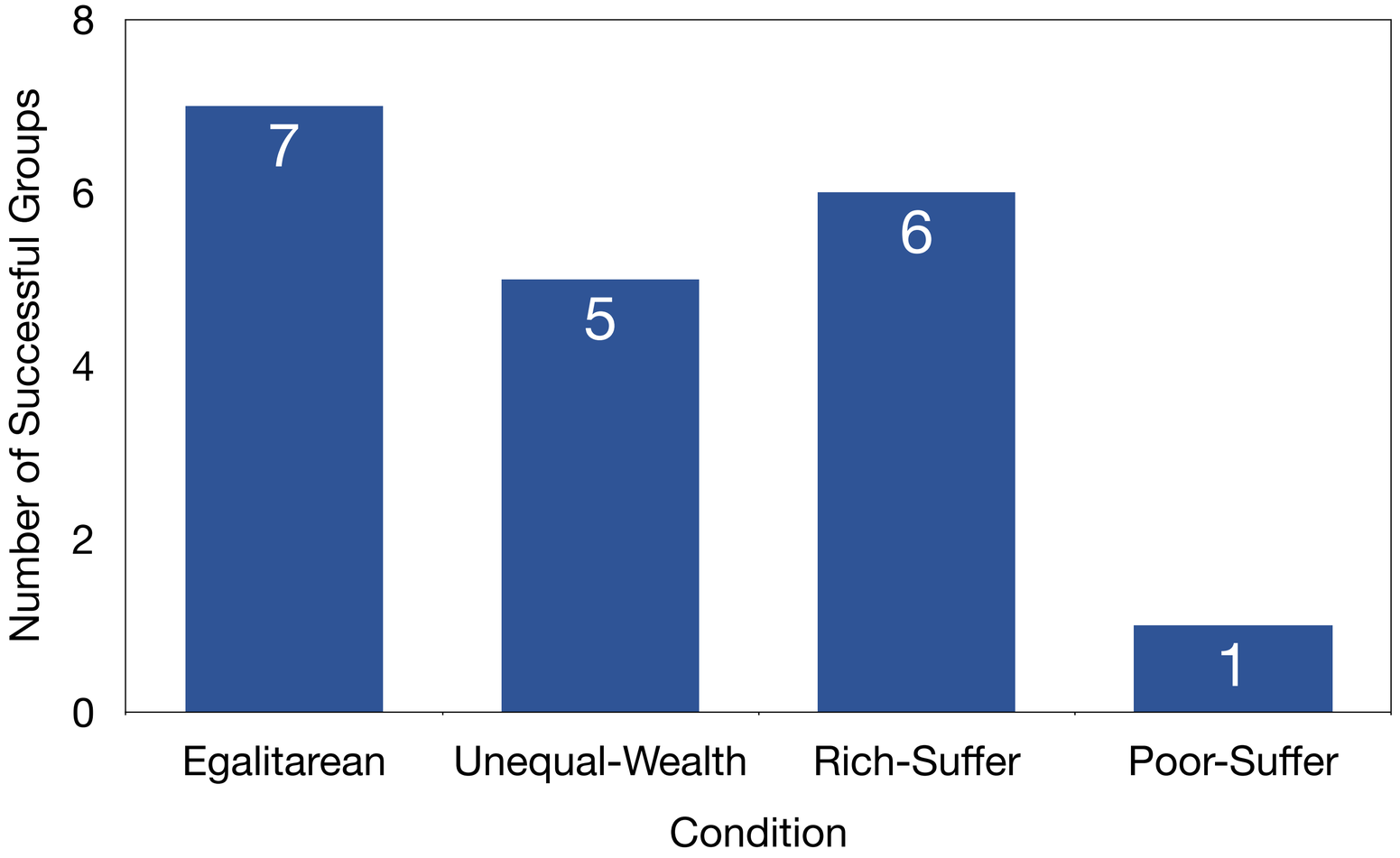
Wealth and Risk Exposure

Burton-Chellew et al. (2013)

- Four conditions:
  - *Egalitarian*: all six group members received an operating fund of €40 faced the same risks if they failed to reach the threshold (e.g., $p = 0.8$)
  - *Unequal-wealth*: group members also faced the same risks, but differed in their resource capacity with two rich players (€80 each) and four poor players (€20 each)
  - *Rich-suffer*: same as unequal-wealth condition, but the risk was higher for rich (e.g., $p = 0.95$) than for poor players (e.g., $p = 0.65$)
  - *Poor-suffer*: same as unequal-wealth condition, but the risk was lower fo rich (e.g., $p = 0.65$) than for poor players (e.g., $p = 0.95$)

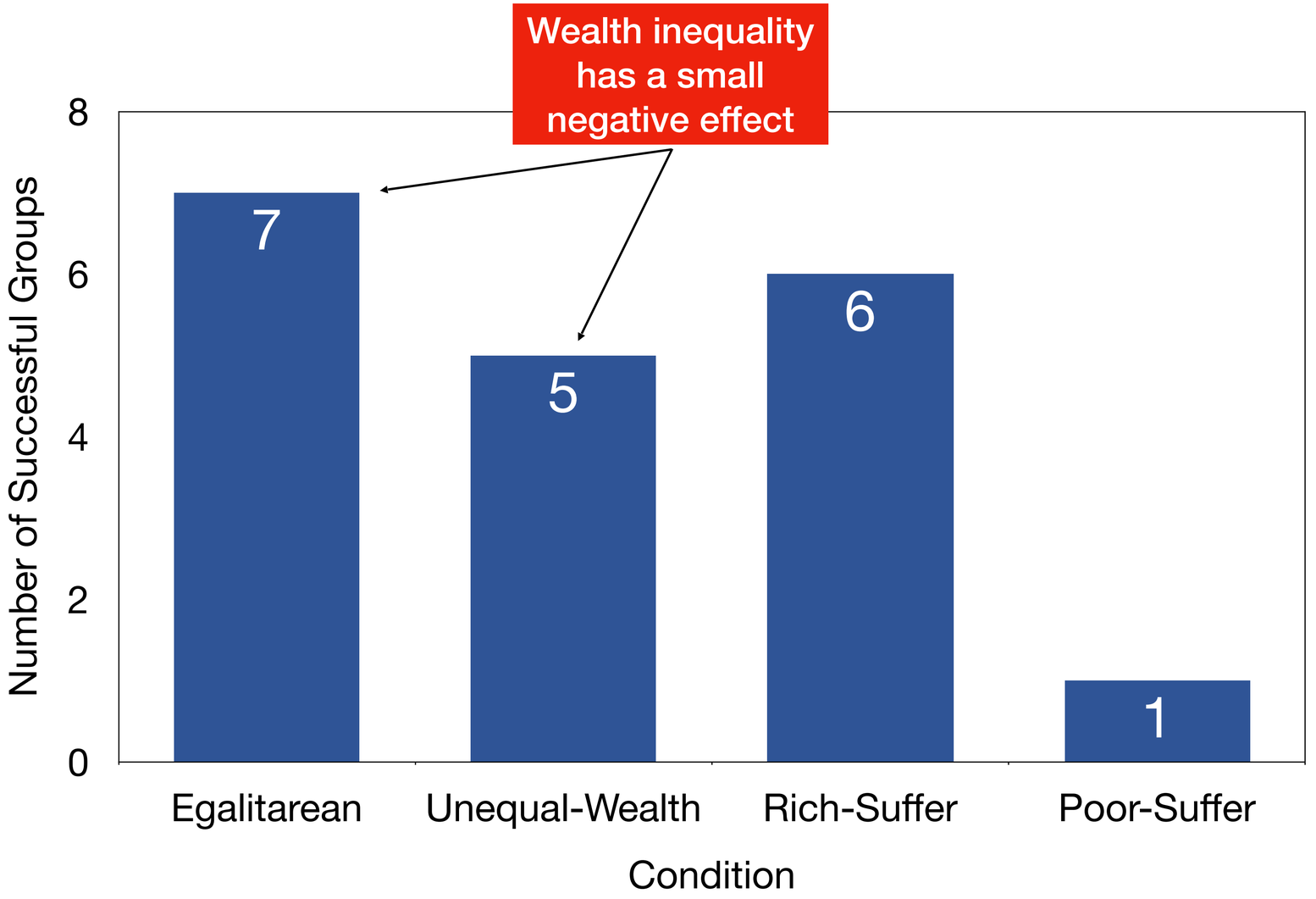
Wealth and Risk Exposure

Burton-Chellew et al. (2013)



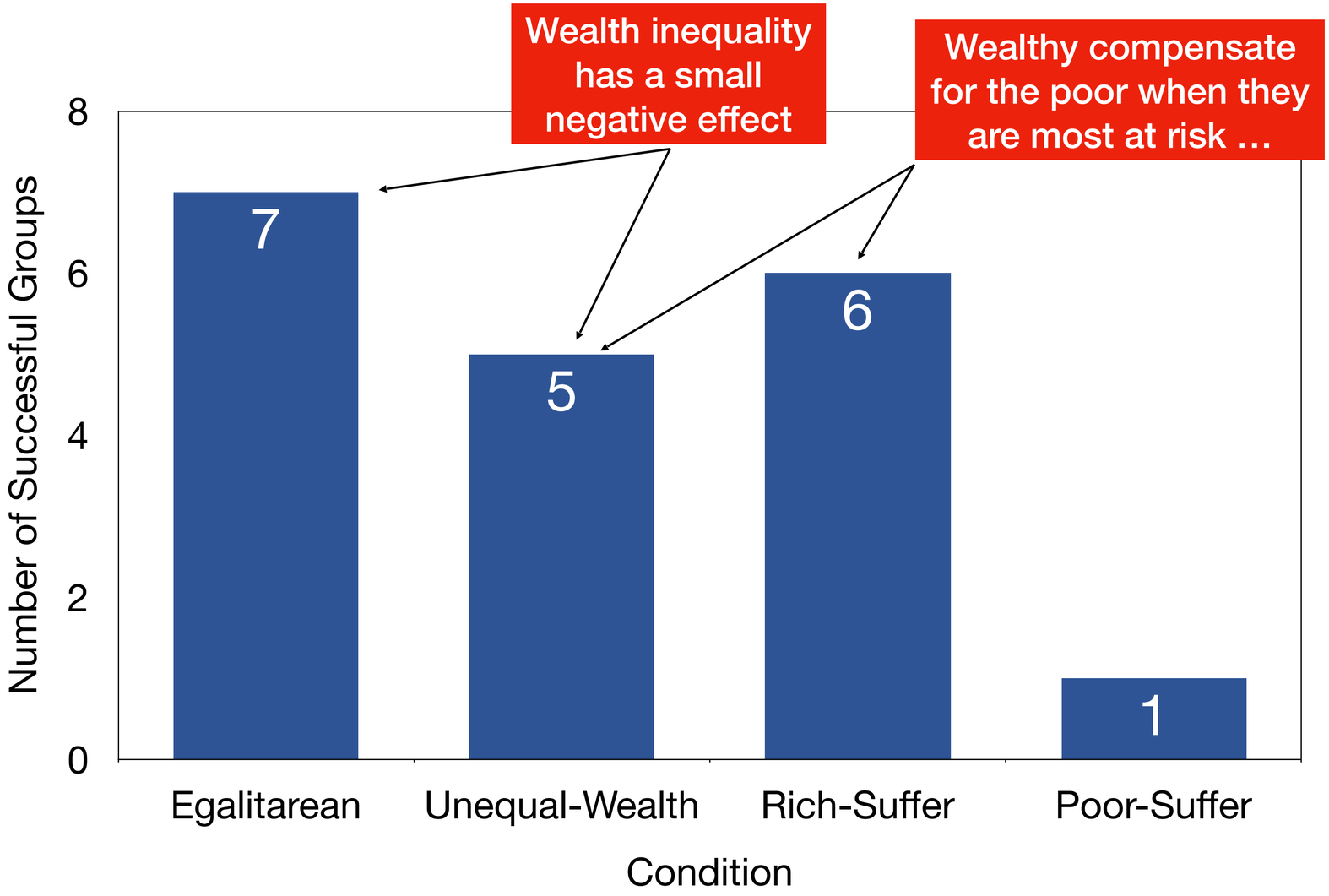
Wealth and Risk Exposure

Burton-Chellew et al. (2013)



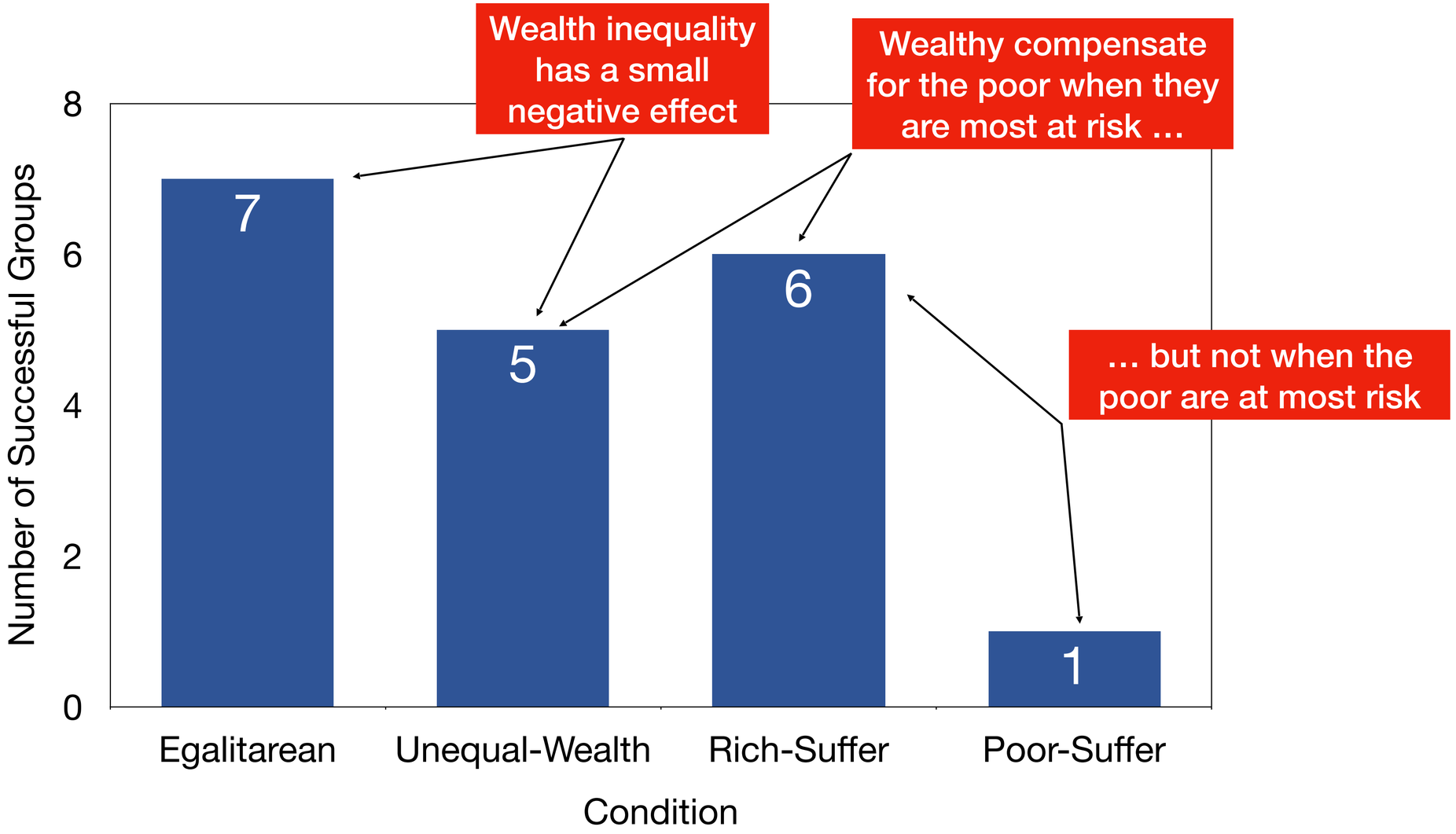
Wealth and Risk Exposure

Burton-Chellew et al. (2013)



Wealth and Risk Exposure

Burton-Chellew et al. (2013)



Implications For Climate Negotiations

- Equity considerations must take centre stage
- Early leadership from powerful countries is a crucial ingredient for collective success
- Effective coordination mechanisms are required to facilitate this process
- The “pledge-and-review” instrument in the Paris Agreement operates in a similar way to the communication instrument used by Tavoni et al. and may help promote equitable burden sharing
- However, the Paris pledges were not based on a common metric—for “pledge-and-review” to be effective, improved methods for assessing and comparing mitigation efforts are needed

Implications For Climate Negotiations

- The results of Burton-Chellew et al. are concerning in light of recent evidence that rich countries perceive climate change as less threatening than poor countries (Lo, 2015; Lo & Chow, 2015)
- This reduced perception of risk is misguided—although the poor will suffer most, this does not mean the rich will not suffer
- A necessary aspect of climate negotiations is to convince rich countries that climate change threatens ruin to all, not just the poor

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Uncertainty About Climate Change

Threshold and Impact Uncertainty

- In the real game of climate change, the location of the dangerous threshold; the effort required to avoid it; and the costs of crossing it are uncertain
- Although the Copenhagen Accord identifies a dangerous temperature threshold of 2°C, other threshold have been identified (e.g., the Paris Agreement introduced a new threshold of 1.5°C)
- Similarly, estimates of the expected damages resulting from catastrophic climate change differ widely
- How do such uncertainties affect the prospects of cooperation?

Uncertainty About Climate Change

Barrett and Dannenberg (2012)

- Variant of collective-risk social dilemma with groups of 10 players
- Players allocated €31: divided into operating fund (€11) and endowment (€20)
  - Operating fund could be used to invest in “weak” or “strong” abatement by purchasing chips (max = 10) at a cost of €0.10 or €1.00, respectively)
  - Endowment could not be spent and was included to ensure players could not become bankrupt
- Game played over a **single round** divided into two stages: (1) communication stage (divided into “proposals” and “pledges”) and (2) contribution stage

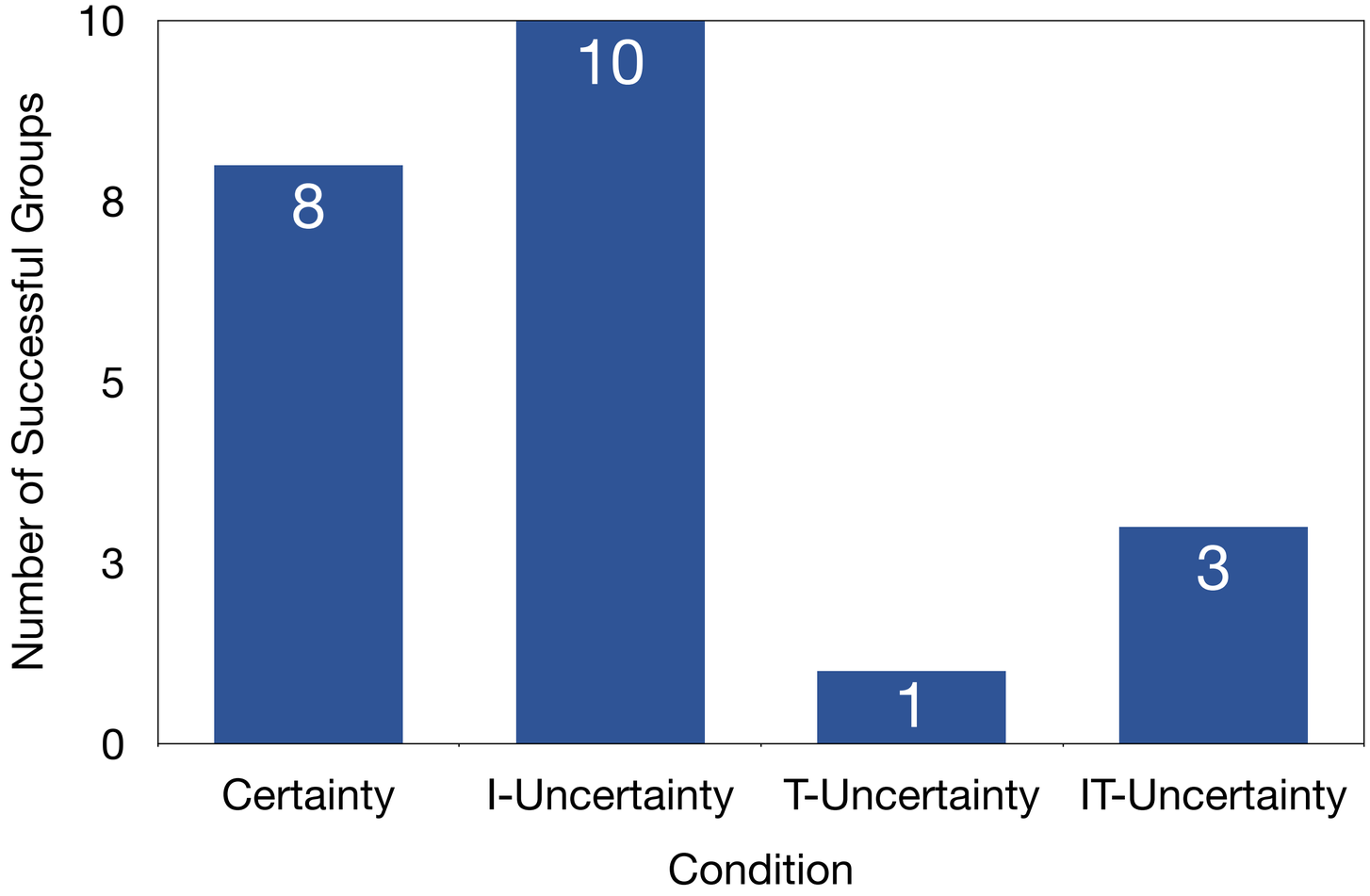
Uncertainty About Climate Change

Barrett and Dannenberg (2012)

- Four conditions:
  - *Certainty*: target sum of investments to be reached was 150 chips, otherwise €15 was deducted from each player's endowment
  - *Impact-uncertainty*: target sum of investments to be reached was 150 chips, otherwise €10—20 was deducted from each player's endowment
  - *Threshold-uncertainty*: target sum of investments to be reached was 100—200 chips, otherwise €15 was deducted from each player's endowment
  - *Impact + threshold uncertainty*: target sum of investments to be reached was 100—200 chips, otherwise €10—20 was deducted from each player's endowment

Uncertainty About Climate Change

Barrett and Dannenberg (2012)



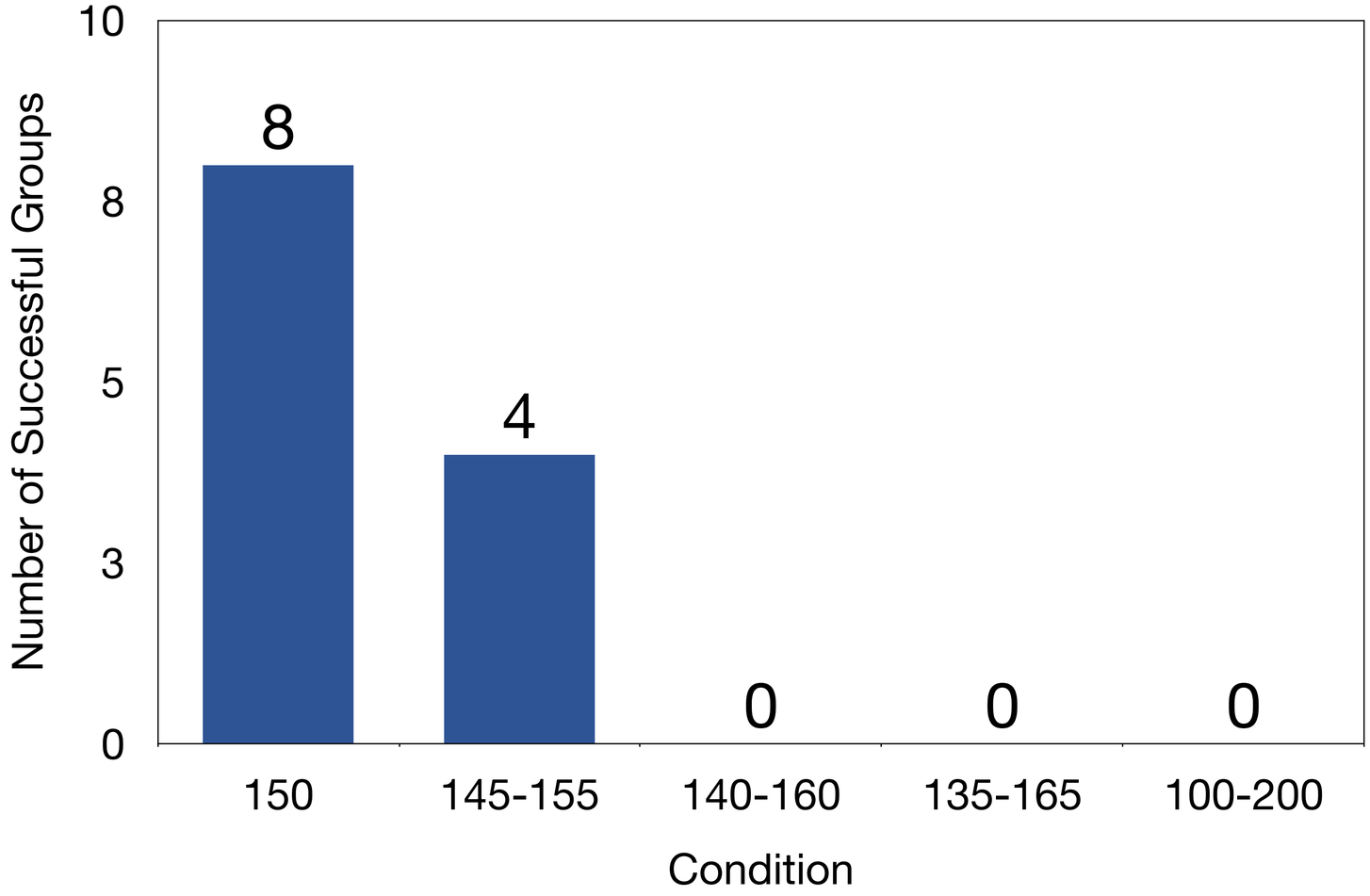
Uncertainty About Climate Change

Barrett and Dannenberg (2014)

- How much must uncertainty about the threshold be reduced?
- *Certainty condition*:
  - Target sum of investments = 150 chips
- *Threshold-uncertainty conditions*:
  - Target sum of investments = 145—155 chips
  - Target sum of investments = 140—160 chips
  - Target sum of investments = 135—165 chips
  - Target sum of investments = 100—200 chips
- Cost deducted from each player's endowment for failing to reach the target was €15 in all conditions

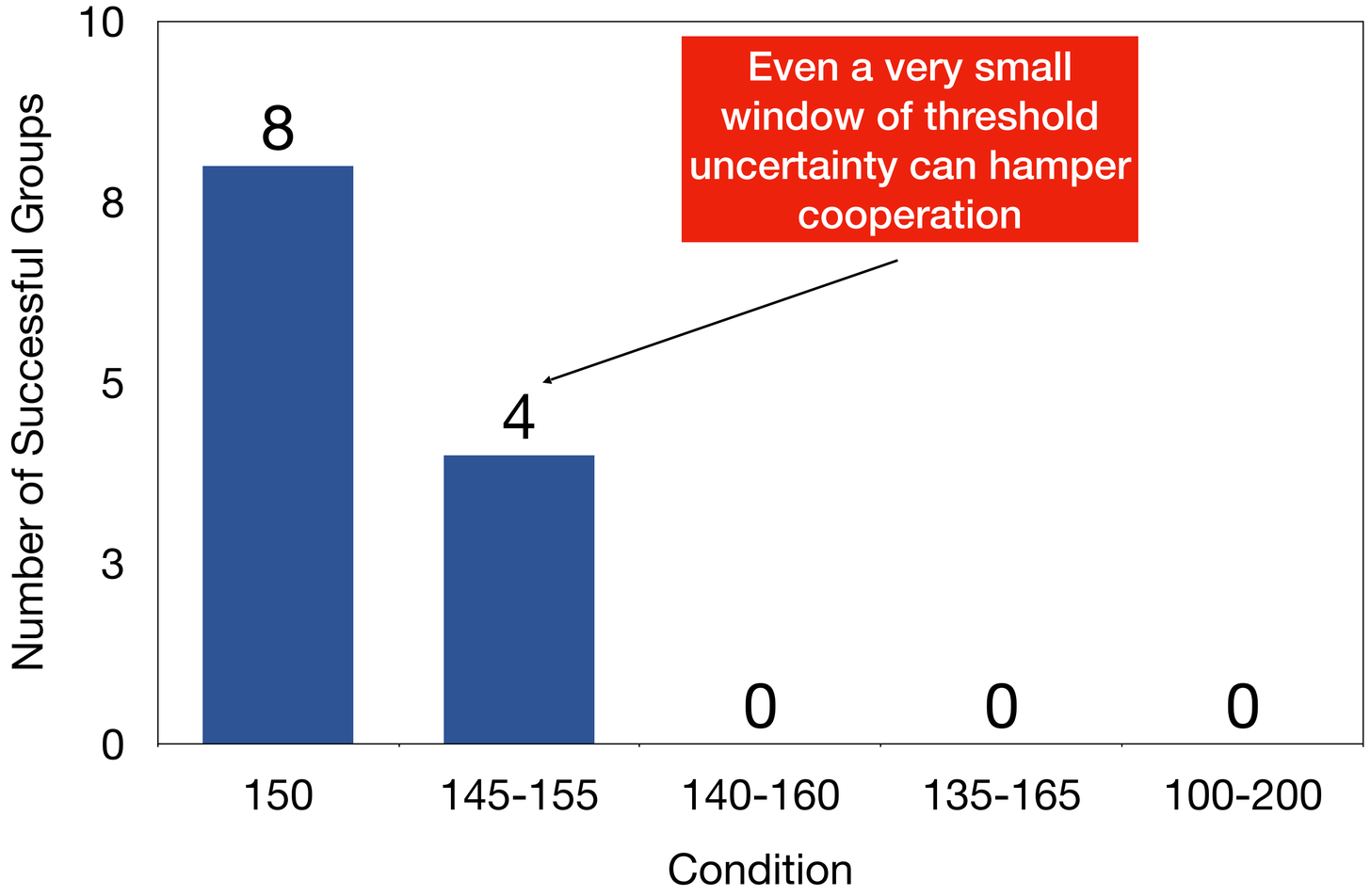
Uncertainty About Climate Change

Barrett and Dannenberg (2014)



Uncertainty About Climate Change

Barrett and Dannenberg (2014)



Uncertainty About Climate Change

Refer to the Introduction of Hurlstone, White, and Newell (2022) for a discussion of these and other studies of threshold uncertainty (located on Moodle)

Implications For Climate Negotiations

- Uncertainty about **climate damages is inconsequential**; uncertainty about the **threshold is crucial**
- Unless the threshold is known with certainty, it is unlikely countries will be able to avert catastrophe
- Threshold uncertainty reduces the credibility of Mother Nature's threat to tip the climate system into chaos if the threshold is breached
- It turns the climate coordination game under a certain threshold into a **prisoners' dilemma**

Implications For Climate Negotiations

- If a **red line** for dangerous climate change could be identified, fear of crossing it would discipline behaviour
- Irreducible uncertainties surrounding the location of the threshold make this prospect unlikely
- Strategic enforcement mechanisms (e.g., trade sanctions) are needed that recreate the conditions that exist when the threshold is certain
- Enforcement mechanisms are **strategic devices**—their purpose is not to be used, but to provide the deterrent necessary to transform behaviour

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Intergenerational Discounting

- Difficulty of avoiding dangerous climate change is exacerbated by climate change's intergenerational nature
- The current generation bears the *costs of cooperation* (*gains of defection*), whilst future generations inherit the *benefits* (*costs)*
- **Temporal discounting** = tendency to prefer immediate over delayed monetary rewards and delayed over immediate costs
  - *Intra-generational discounting*: temporal discounting over short-term time horizons
  - *Inter-generational discounting*: temporal discounting over extremely long-term time horizons

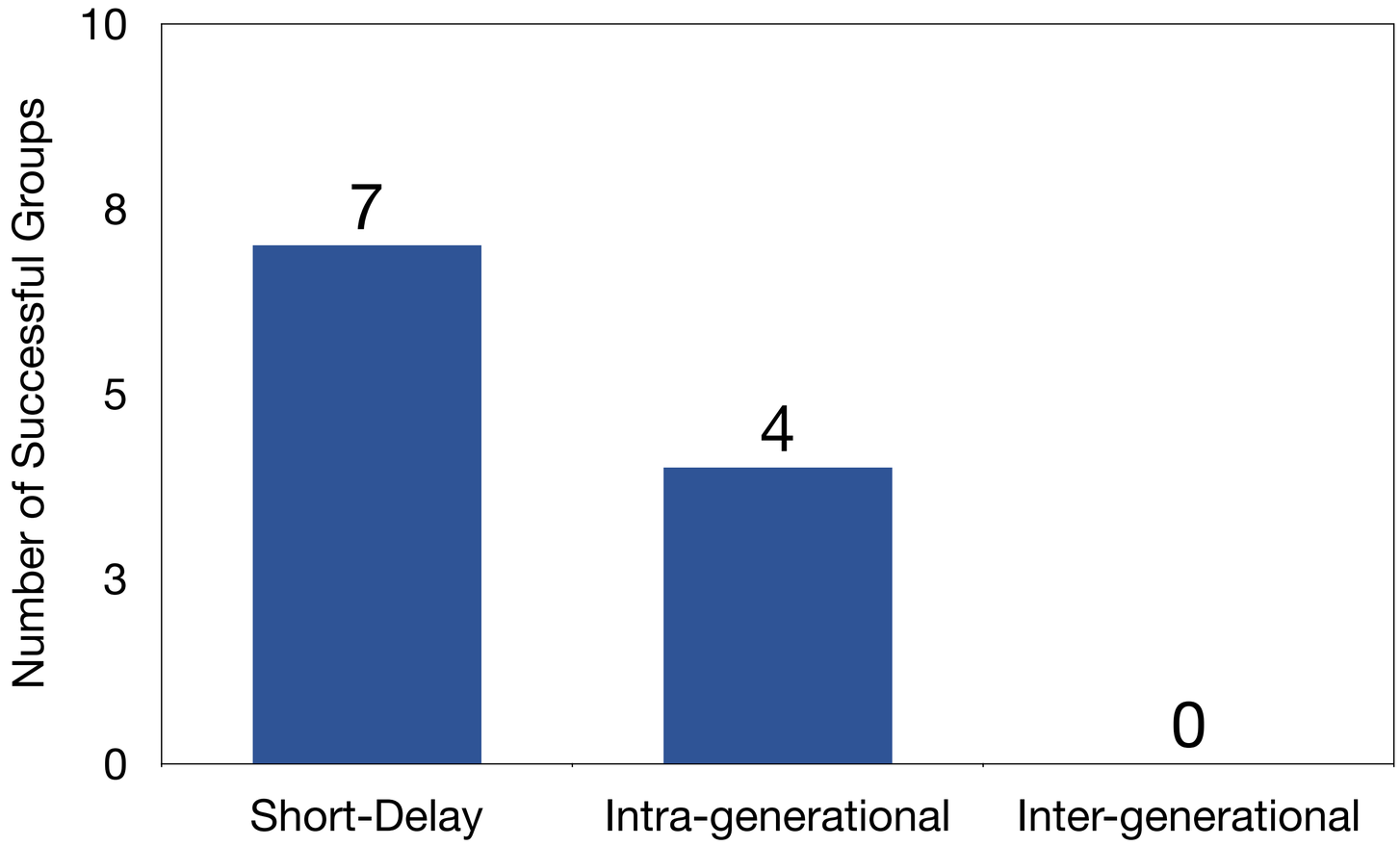
Intergenerational Discounting

Jacquet et al. (2013)

- Collective-risk social dilemma in which players were given an operating fund of €40 and an endowment of €45
- Players always received the left-overs of their operating fund at the end of the game, whereas the endowment was only awarded if catastrophe was avoided
- Three conditions:
  - *Short-delay*: endowment delayed by 1 day
  - *Intra-generational*: endowment delayed by 7 weeks
  - *Inter-generational*: endowment delayed by several decades with wider range of beneficiaries

Intergenerational Discounting

Jacquet et al. (2013)



Intergenerational Discounting

Jacquet et al. (2013)

- Intergenerational discounting is a major impediment to cooperation—immediate monetary rewards matter most
- Paints a sobering picture for climate negotiations
- Due to intergenerational discounting, cooperation will be greatly undermined if short-term gains can arise only from defection
- Suggests the necessity of introducing powerful short-term incentives to cooperate such as punishment, reward, or **reputation**

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

Factors Affecting Cooperation

Facilitators and Impediments

1. Perception of risk
2. Inequality
3. Uncertainty about dangerous climate change
4. Intergenerational discounting
5. Reputation

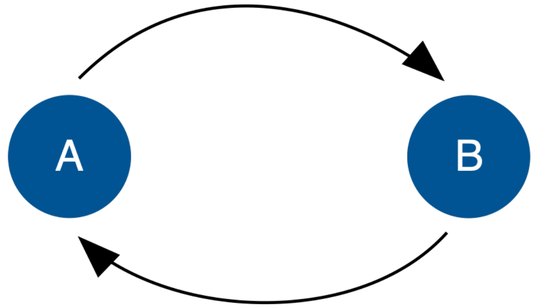
Reputation

Milinski et al. (2006)

- In the experiments considered so far, the reputation of players cannot be used as a tool to sustain cooperation
- However, if a **good reputation** can be used as a **currency** by which to **obtain rewards** and **avoid punishments**, stable cooperation can be achieved
- Milinski et al. (2006) alternated rounds of a “climate cooperation” game with rounds of an “indirect reciprocity” game

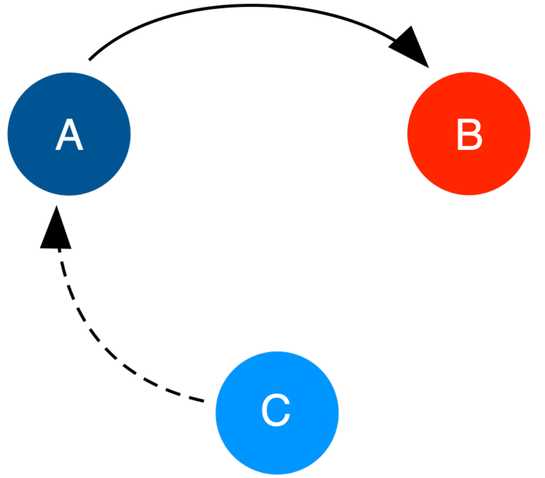
The Language of Reciprocity

a Direct Reciprocity



"You scratch my back, and I'll scratch yours"

b Indirect Reciprocity



"I scratch your back, and someone else will scratch mine"

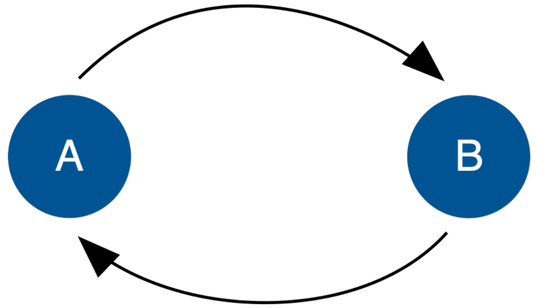
Reputation

Milinski et al. (2006)

- In the experiments considered so far, the reputation of players cannot be used as a tool to sustain cooperation
- However, if a **good reputation** can be used as a **currency** by which to **obtain rewards** and **avoid punishments**, stable cooperation can be achieved
- Milinski et al. (2006) alternated rounds of a “climate cooperation” game with rounds of an “indirect reciprocity” game

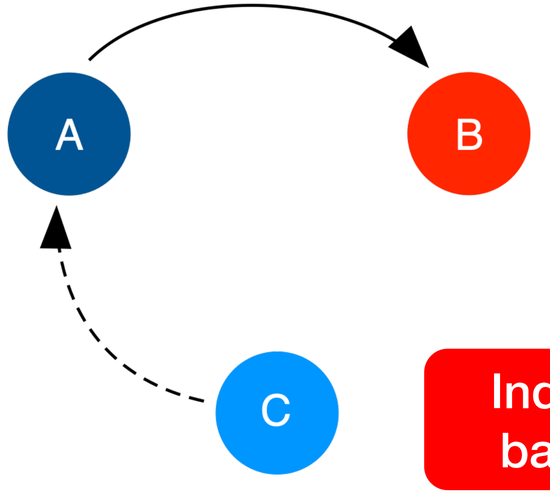
The Language of Reciprocity

a Direct Reciprocity



"You scratch my back, and I'll scratch yours"

b Indirect Reciprocity



"I scratch your back, and someone else will scratch mine"

Indirect reciprocity is based on reputation

Reputation

Milinski et al. (2006)

- **Climate cooperation game:**
  - Groups of six players given €12 operating fund and could invest €0, €1, or €2 into a climate account over 10 rounds
  - Any money invested in the climate account was doubled in value and used to fund a press advertisement on climate protection
- **Indirect reciprocity game:**
  - On each round, players adopted the role of “donor” or “receiver” once
  - When given the role of donor, a player had to decide whether to give a reward of €3 to another player at a cost of €1.50 to themselves
  - When given the role of receiver, a player could potentially receive a €3 reward from another player

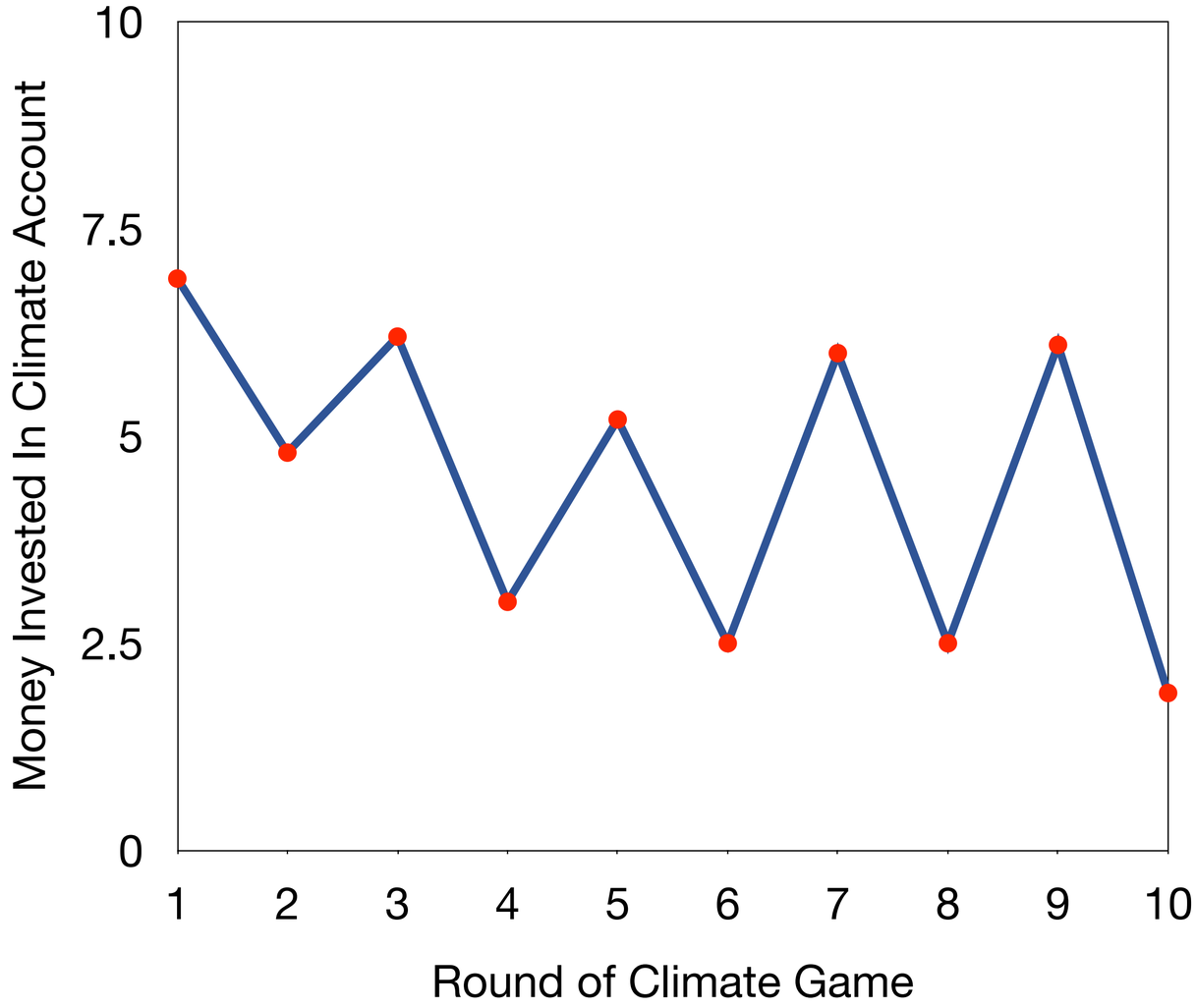
Reputation

Milinski et al. (2006)

- The key manipulation was that on “odd” rounds of the climate cooperation game, investments in the climate account were made public, whereas on “even” rounds they were made anonymous
- Thus, on **public rounds,** there was an incentive for players to cooperate by investing in the climate account
- This is because **cooperation affords a positive reputation**, which should be **rewarded** in the indirect reciprocity game, whereas **defection affords a negative reputation**, which may incur **punishment**
- On **anonymous rounds**, the incentive to cooperate and the disincentive to free ride is removed because reputation cannot be used to obtain benefits in the indirect reciprocity game

Reputation

Milinski et al. (2006)



Implications For Climate Negotiations

- The Milinski et al. (2006) experiment provides a “proof of concept” for **climate clubs** (Nordhaus, 2015)
- A small and powerful “coalition of the willing” establishes a voluntary group that produces **economic or non-economic benefits**
- These “club goods” could be, for example, development of a new technology, pooled finances, or common technological standards
- Joining the club is contingent on having a **good reputation for climate protection**, as is sustaining club membership

An Experimental Assessment of The Paris Agreement

An Experimental Assessment of The Paris Agreement

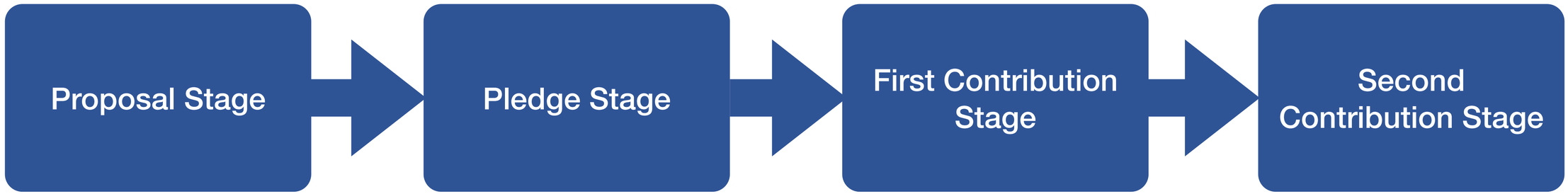
Testing Pledge-And-Review

- The key feature of the Paris Agreement is its “pledge-and-review” enforcement mechanism (“soft power”)
- Provides countries with the opportunity to express their approval or disapproval of other countries’ pledges and contributions
- Designed to promote **praising of role models** and **naming and shaming** of free riders
- Will this mechanism ensure that the Paris Agreement goals are met?

An Experimental Assessment of The Paris Agreement

Barrett and Dannenberg (2016)

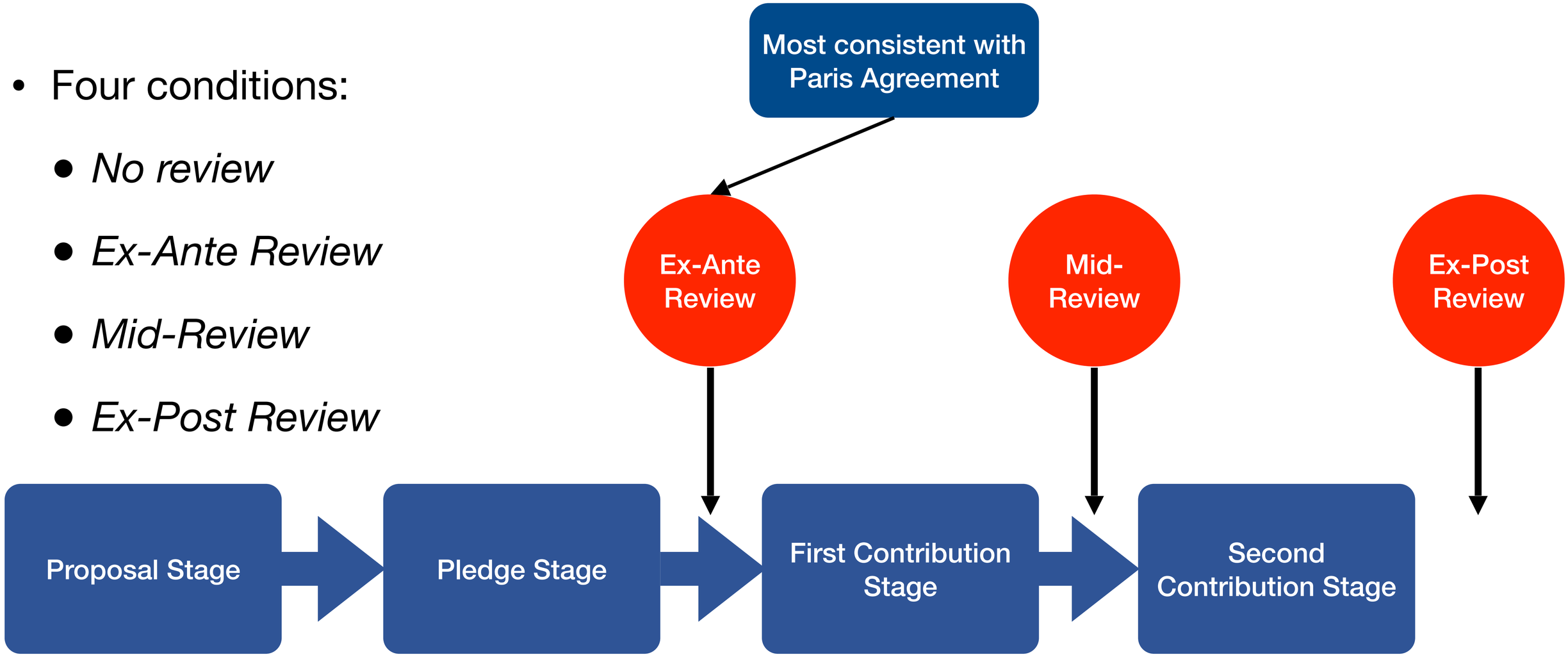
- Groups of 5 players
- Uncertainty about the dangerous threshold
- One shot-game played over a sequence of four stages:



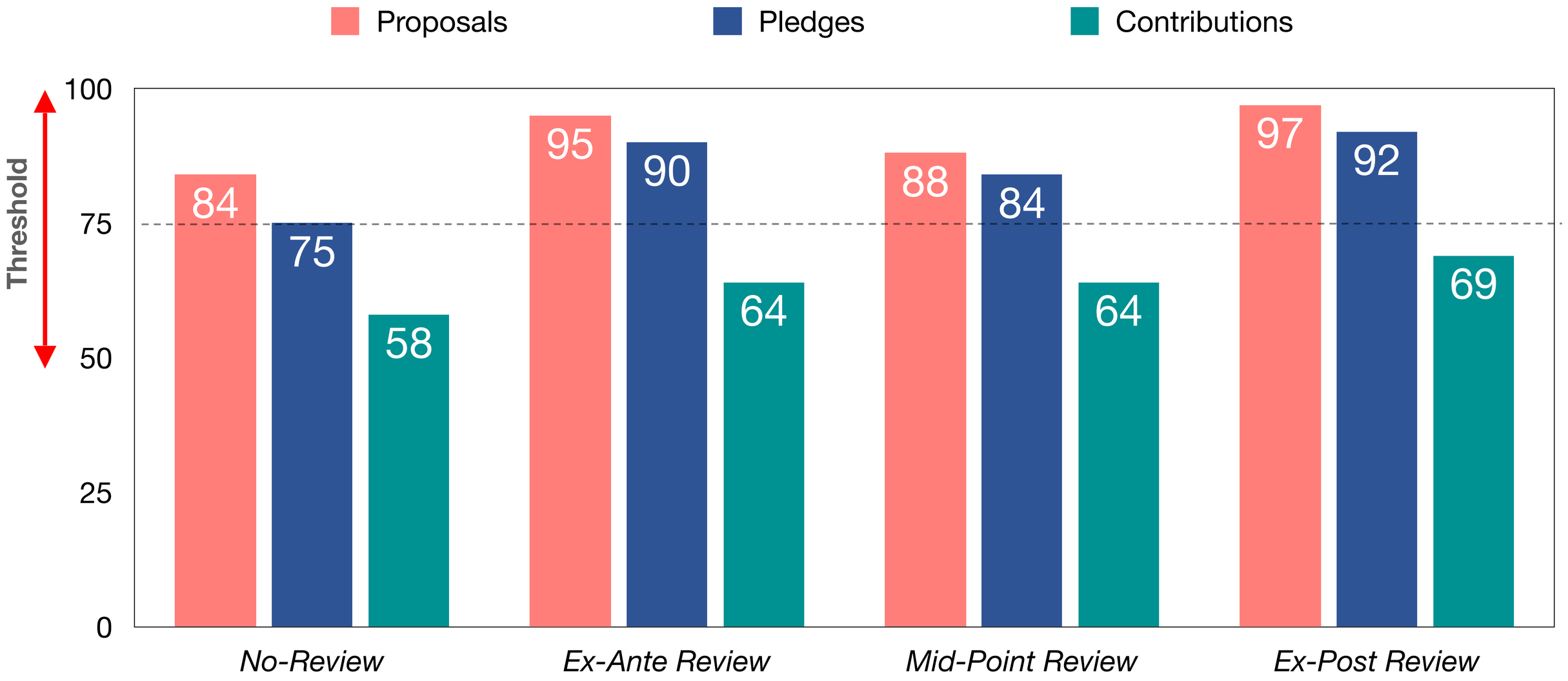
An Experimental Assessment of The Paris Agreement

Barrett and Dannenberg (2016)

- Four conditions:
  - *No review*
  - *Ex-Ante Review*
  - *Mid-Review*
  - *Ex-Post Review*



An Experimental Assessment of The Paris Agreement



Implications For Climate Negotiations

- The Paris pledges will at best limit warming to **2.7°C**, so as in Barrett and Dannenberg's (2016) experiment countries have already **pledged to do less** than is required to reach the **collective target**
- Their results suggest that even with a pledge-and-review mechanism, **countries' actual emission reductions** are likely to be **less than their pledges**
- Regrettably, countries **post-Paris emission reductions bear this prediction out**
- Suggests **complimentary strategies** running in parallel with the Paris Agreement are required to achieve the level of cooperation needed to fulfil its objective of limiting warming to 2°C

Implications For Climate Negotiations

- An effective strategy is to exploit **linkage mechanisms** that co-opt non-climate institutions for international cooperation to protect the climate
- The **Montreal Protocol on Substances that Deplete the Ozone Layer** has served an important function in protecting the climate
- This is because many of the ozone-depleting substances phased out under this agreement are also greenhouse gases that contribute to climate change
- Prior to the Paris negotiations (Nov, 2015), **The Kigali Amendment** to the Montreal Protocol was negotiated to phase down the use of global-warming-inducing Hydrofluorocarbons

Conclusions

Implications For Climate Negotiations

1. Countries must be convinced of the very **high risks** of dangerous climate change
2. Greater **transparency** and **improved methods for determining comparability of effort** are needed to promote equitable burden sharing
3. The handicap of threshold uncertainty and ineffectiveness of pledge-and-review suggests **strategic enforcement mechanisms** must be incorporated into the Paris Agreement
4. Before then, climate negotiators can **co-opt existing non-climate institutions** for international cooperation that already contain such mechanisms to achieve climate benefits
5. The intergenerational nature of climate change is a significant impediment to cooperation, but this impediment might be overcome through the formation of **climate clubs** that yield short-term benefits that exceed the immediate rewards of free riding